NORTH-HOLLAND

Canonical Form of Symplectic Matrix Pencils

Augusto Ferrante*
Department of Electrical, Mechanical, and Management Engineering
University of Udine
via delle Scienze, 208
33100 Udine, Italy

and

Bernard C. Levy†
Department of Electrical and Computer Engineering
University of California
Davis, California 95616

Submitted by Paul A. Fuhrmann

ABSTRACT

We obtain a real canonical form for real symplectic pencils. It extends earlier results which were derived over the complex field, so that the canonical form was complex, and which were limited to the case where the elementary divisors of eigenvalues on the unit circle have an even degree. © 1998 Elsevier Science Inc.

1. INTRODUCTION

Let

$$K = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}. \tag{1.1}$$

* E-mail: augusto@picolit.diegm.uniud.it.The research of this author was performed in part at the Institute of Theoretical Dynamics, University of California at Davis, with support provided by a CNR fellowship.
† E-mail: levy@ece.ucdavis.edu.

LINEAR ALGEBRA AND ITS APPLICATIONS 274:259–300 (1998)

655 Avenue of the Americas, New York, NY 10010

0024-3795/98/$19.00
PII S0024-3795(97)00347-9

Two real matrices H and S are said to be Hamiltonian and symplectic, respectively, if $H^T K + KH = 0$ and $S^T KS = K$. Observe that if H is Hamiltonian, then $S_1 = e^H$, $S_2 = (H + I_n)(H - I_n)^{-1}$, and $S_3 = (H - I_n)(H + I_n)^{-1}$ are all symplectic matrices, where S_2 and S_3 are defined only if H does not have eigenvalues at 1 or -1, respectively. Note also that the Cayley transformation used to generate S_2 and S_3 can be inverted and used to convert a symplectic matrix S into a Hamiltonian one, as long as S does not have eigenvalues either at 1 or at -1.

Starting with the work of Williamson [17–19], the symmetries of Hamiltonian and symplectic matrices have been exploited by a number of researchers, such as Cikunov [5, 4], Ciampi [3, 2], Laub and Meyer [13], Burgoyne and Cushman [1], Djokovic et al. [6], and Lancaster and Rodman [12], to obtain canonical forms for these two classes of matrices. The main difficulty is, for the Hamiltonian case, to characterize the eigenstructure corresponding to eigenvalues on the imaginary axis or at the origin, and for the symplectic case, to characterize the eigenstructure corresponding to eigenvalues on the unit circle or at ± 1. Hamiltonian and symplectic matrices arise naturally in a wide variety of problems of mechanics, optimal control [12], and signal processing, so that studies of the internal structure of such matrices have a wide applicability. For example, the canonical form of Hamiltonian matrices was employed recently in [14] to perform a complete classification of multivariate stationary Gaussian reciprocal diffusions.

A partial study of canonical forms of symplectic pencils of the form $sE - tA$, with E and A not necessarily invertible, was undertaken by Wimmer [20] in the context of an analysis of the discrete-time algebraic Riccati equation. However, Wimmer's results are restricted to the case where the elementary divisors corresponding to eigenvalues on the unit circle have an even degree. Moreover, the analysis of [20] was performed over the complex field, and the canonical form obtained is, in general, complex even for real pencils. The objective of this paper is to extend Wimmer's results by constructing a real canonical form for arbitrary real symplectic pencils. This generalization is needed to extend the results of [14] to discrete-time stationary Gaussian reciprocal processes.

The approach we follow exploits the symmetries of symplectic pencils to restructure the Weierstrass canonical form of regular matrix pencils. This form, which is described in detail in [7, Chapter 12], is the analog for matrix pencils of the Jordan form of a matrix. Note that this canonical form is often attributed to Kronecker, although according to [16, Chapter 9], it was first derived by Weierstrass for regular pencils, and subsequently extended to singular pencils by Kronecker. The canonical form of symplectic pencils is described in Section 2, and after deriving some preliminary results in Section 3, its construction is given in Section 4.

1.1. *Mathematical Setting and Definitions*

Consider a matrix pencil

$$P(s,t) = sE - tA, \tag{1.2}$$

where E and A are two real matrices with the same size. The pencil is said to be square if E and A are square matrices, and regular if its determinant does not vanish identically.

An *eigenvalue* of the pencil $P(s,t)$ is a number $z_0 \in \mathbb{C} \cup \{\infty\}$ such that there exists a pair $(s_0, t_0) \in \mathbb{C}^2 \setminus \{(0,0)\}$ for which $z_0 = s_0/t_0$ and $\det P(s_0, t_0) = 0$, with the convention that $s_0/0 = \infty$ for any complex s_0. To characterize the properties of matrix pencils which remain invariant under coordinate changes, the following equivalence relation may be introduced.

DEFINITION 1.1. Two real matrix pencils $P_1(s,t)$ and $P_2(s,t)$ are *equivalent*, which we denote by

$$P_1(s,t) \sim P_2(s,t), \tag{1.3}$$

if there exist two real invertible matrices V and W such that

$$P_1(s,t)V = WP_2(s,t). \tag{1.4}$$

The following definition introduces symplectic pencils and symplectic pairs.

DEFINITION 1.2. A pair $(P(s,t), K)$ formed by a regular real matrix pencil $P(s,t) = sE - tA$ of size $2n \times 2n$ and a real $2n \times 2n$ nonsingular skew-symmetric matrix K is said to be *symplectic* if

$$E^T KE = A^T KA. \tag{1.5}$$

If K is given by the right-hand side of (1.1), the pencil $P(s,t)$ itself is said to be symplectic.

Recall that for a real pencil $P(s,t)$, if z_0 is an eigenvalue of P, so is its complex conjugate z_0^*. The symplectic property (1.5) imposes additional symmetries to the eigenstructure of $P(s,t)$, which are described in Section 3.

We now define an equivalence relation between symplectic pairs. The goal of our paper is to construct a canonical form under this equivalence relation.

DEFINITION 1.3. Two symplectic pairs $(P_1(s,t), K_1)$ and $(P_2(s,t), K_2)$ are said to be *symplectically equivalent*, which we denote by

$$(P_1(s,t), K_1) \overset{S}{\sim} (P_2(s,t), K_2), \tag{1.6}$$

if there exist two real invertible matrices V and W such that

$$P_1(s,t)V = WP_2(s,t), \tag{1.7a}$$

$$W^T K_1 W = K_2. \tag{1.7b}$$

In 1867, Weierstrass showed that under the equivalence relation (1.3), a regular pencil can be brought to a canonical form with the structure shown in Theorem 3.1 below. The purpose of the present paper is to reorganize the Weierstrass canonical form to exploit the symmetries of the class of symplectic pairs.

Before describing the main result, we introduce some notation that will be useful in the sequel. First recall that, given two matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times q}$, and denoting by a_{ij} the (i,j)th element of the matrix A, the *Kornecker product* $A \otimes B$ of A and B takes the form

$$A \otimes B := \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{bmatrix}. \tag{1.8}$$

It is therefore a block matrix of size $np \times mq$ which admits $a_{ij}B$ as its (i,j)th block of size $p \times q$. The reader is referred to [10] for a discussion of the properties of the matrix Kronecker product.

Given two matrices A and B, we write $A \oplus B$ the block-diagonal matrix

$$A \oplus B := \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}. \tag{1.9}$$

We denote by Z_r and Σ_r and $r \times r$ matrices given respectively by

$$Z_r = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{1.10}$$

and

$$\Sigma_r = \begin{bmatrix} 0 & \cdots & 0 & 0 & (-1)^{r-1} \\ 0 & \cdots & 0 & (-1)^{r-2} & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{1.11}$$

For a complex number $a = \sigma + j\omega$, $J_r(a)$ denotes a Jordan block of size r with eigenvalue a, i.e.

$$J_r(a) = aI_r + Z_r. \tag{1.12}$$

Similarly, $J_{2r}(a, a^*)$ represents the $2r \times 2r$ real Jordan block obtained by pairing the complex Jordan blocks of size r associated to a and a^*:

$$J_{2r}(a, a^*) = I_r \otimes \begin{bmatrix} \sigma & -\omega \\ \omega & \sigma \end{bmatrix} + Z_r \otimes I_2. \tag{1.13}$$

2. MAIN RESULT

The canonical form that we propose for symplectic pencils has the following structure.

THEOREM 2.1. *Under the symplectic equivalence relation* (1.6), *a symplectic pair* $(P(s,t), K)$ *can be transformed to a canonical form* $(\tilde{P}(t,s), \tilde{K})$, *where*

$$\tilde{P}(s,t) = \bigoplus_{i=1}^{l} (sE_i - tA_i), \qquad \tilde{K} = \bigoplus_{i=1}^{l} K_i, \tag{2.1}$$

and the blocks E_i, A_i, and K_i are of four possible types:

Type 1. *The blocks E_i, A_i, and K_i corresponding to a real eigenvalue pair (a_i, a_i^{-1}), with $|a_i| < 1$, take the form*

$$E_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} I_{r_k} & 0 \\ 0 & J_{r_k}^T(a_i) \end{bmatrix}, \tag{2.2a}$$

$$A_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} J_{r_k}(a_i) & 0 \\ 0 & I_{r_k} \end{bmatrix}, \tag{2.2b}$$

$$K_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} 0 & -I_{r_k} \\ I_{r_k} & 0 \end{bmatrix}. \tag{2.2c}$$

Type 2. *The blocks E_i, A_i, and K_i corresponding to a complex eigenvalue quadruple $(a_i, a_i^*, a_i^{-1}, a_i^{-*})$, such that $a_i = \sigma_i + j\omega_i$ with $\omega_i > 0$ and $|a_i| < 1$, admit the structure*

$$E_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} I_{2r_k} & 0 \\ 0 & J_{2r_k}^T(a_i, a_i^*) \end{bmatrix}, \tag{2.3a}$$

$$A_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} J_{2r_k}(a_i, a_i^*) & 0 \\ 0 & I_{2r_k} \end{bmatrix}, \tag{2.3b}$$

$$K_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} 0 & -I_{2r_k} \\ I_{2r_k} & 0 \end{bmatrix}. \tag{2.3c}$$

Type 3. *The blocks E_i, A_i, and K_i corresponding to a complex eigenvalue pair $(e^{j\theta_i}, e^{-j\theta_i})$ on the unit circle, with $0 < \theta_i < \pi$, admit the structure*

$$E_i = \bigoplus_{k=1}^{p_i} \left(I_{2r_k} - J_{2r_k}(jb_i\ -jb_i)\right), \tag{2.4a}$$

$$A_i = \bigoplus_{k=1}^{p_i} \left(I_{2r_k} + J_{2r_k}(jb_i\ -jb_i)\right), \tag{2.4b}$$

$$K_i = \bigoplus_{k=1}^{p_i} \kappa_k\left(\Sigma_{r_k} \otimes \Sigma_2^{r_k}\right) \tag{2.4c}$$

with $\kappa_k = \pm 1$ *and* $b_i = \tan(\theta_i/2)$.

Type 4. *The blocks E_i, A_i, and K_i corresponding to eigenvalues located at $\epsilon_i = \pm 1$ take the form*

$$E_i = \bigoplus_{k=1}^{p_{ei}} \left(I_{2r_k} - Z_{2r_k}\right) \oplus \left(\bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} I_{2r_k'+1} - Z_{2r_k'+1} & 0 \\ 0 & \left(I_{2r_k'+1} + Z_{2r_k'+1}\right)^T \end{bmatrix}\right), \tag{2.5a}$$

$$A_i = \epsilon_i\left\{\bigoplus_{k=1}^{p_{ei}} \left(I_{2r_k} + Z_{2r_k}\right) \oplus \left(\bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} I_{2r_k'+1} + Z_{2r_k'+1} & 0 \\ 0 & \left(I_{2r_k'+1} - Z_{2r_k'+1}\right)^T \end{bmatrix}\right)\right\}, \tag{2.5b}$$

$$K_i = \left(\bigoplus_{k=1}^{p_{ei}} \kappa_k \Sigma_{2r_k}\right) \oplus \left(\bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} 0 & -I_{2r_k'+1} \\ I_{2r_k'+1} & 0 \end{bmatrix}\right) \tag{2.5c}$$

with $\kappa_k = \pm 1$.

REMARKS.

(1) Each block K_i is clearly skew-symmetric of even size, and each pair $(P_i(s,t) = sE_i - tA_i, K_i)$ is symplectic. This implies that $(\tilde{P}(s,t), \tilde{K})$ is a symplectic pair.

(2) The blocks K_i are orthogonal:

$$K_i K_i^T = -K_i^2 = I, \qquad i = 1, 2, \dots, l. \tag{2.6}$$

(3) An interesting feature of the triples (E_i, A_i, K_i) is that they admit the symmetry

$$A_i^T = \epsilon_i K_i E_i K_i^T, \tag{2.7}$$

with $\epsilon_i = 1$ for blocks of type 1, 2, or 3, and $\epsilon_i = \pm 1$ for blocks of type 4, depending on whether they are associated to eigenvalues at 1 or -1.

(4) In Theorem 2.1 we have regrouped the pairs (λ, λ^{-1}) of reciprocal eigenvalues. If E is singular, the pencil $P(s, t)$ admits $(0, \infty)$ as an eigenvalue pair. This pair is included among the blocks of type 1.

(5) The canonical form $(\tilde{P}(s, t), \tilde{K})$ is defined only up to an ordering of the diagonal blocks $(P_i(s, t), K_i)$. Such an ordering is uninteresting and is therefore not specified.

(6) The skew-symmetric blocks K_i employed in the above decomposition have the same structure as those used for Hamiltonian matrices in [6]. Since we are dealing with real pencils, they do not coincide with the blocks employed in [20], where the complex case was considered.

(7) The signs $\kappa_k = \pm 1$ appearing in the blocks of type 3 and 4 are invariants of the symplectic pair $(P(s, t), K)$ under the equivalence relation (1.6), and are of the same type as the *inertial invariants* of [13, 20], or the *sign characteristic* of [8].

3. PRELIMINARY RESULTS

Our derivation of Theorem 2.1 makes use of the Weierstrass canonical form of an arbitrary regular pencil, which has the following structure [7, Chapter 12].

THEOREM 3.1. *Under the equivalence relation* (1.2) *a regular real matrix pencil can be brought to the block-diagonal form*

$$s\mathscr{E}_D - t\mathscr{A}_D := \bigoplus_{i=1}^{l} (s\mathscr{E}_i - t\mathscr{A}_i), \tag{3.1}$$

where the blocks $\mathscr{E}_i$ and A_i are of three possible types:

Type R. *To each distinct real finite eigenvalue a_i corresponds a block of the form*

$$\mathscr{E}_i = I = \bigoplus_{k=1}^{p_i} I_{r_k}, \tag{3.2a}$$

$$\mathscr{A}_i = \bigoplus_{k=1}^{p_i} J_{r_k}(a_i). \tag{3.2b}$$

Type C. *To each pair (a_i, a_i^*) of finite complex eigenvalues corresponds a block with the structure*

$$\mathscr{E}_i = I = \bigoplus_{k=1}^{p_i} I_{2r_k}, \tag{3.3a}$$

$$\mathscr{A}_i = \bigoplus_{k=1}^{p_i} J_{2r_k}(a_i, a_i^*). \tag{3.3b}$$

Type I. *To an eigenvalue at infinity, if present, corresponds a block of the form*

$$\mathscr{E}_i = \bigoplus_{k=1}^{p_i} J_{r_k}(0) = \bigoplus_{k=1}^{p_i} Z_{r_k}, \tag{3.4a}$$

$$\mathscr{A}_i = I = \bigoplus_{k=1}^{p_i} I_{r_k}. \tag{3.4b}$$

The following results will also be of use in the derivation of Theorem 2.1.

LEMMA 3.1. *If $(sE - tA, K)$ is a symplectic pair, the pencil $sE - tA$ admits the equivalence relation*

$$sE - tA \sim (tE - sA)^T. \tag{3.5}$$

Proof. From the symplectic property (1.5) it follows that, for arbitrary α, β, s, and t,

$$(\alpha E - \beta A)^T K(sE - tA) = (tE - sA)^T K(\beta E - \alpha A). \tag{3.6}$$

Because of the regularity of the pencil $sE - tA$, we can choose α and β real such that both $(\alpha E - \beta A)^T$ and $\beta E - \alpha A$ are nonsingular (i.e., such that neither α/β nor β/α are eigenvalues of the pencil $sE - tA$). Thus, both $(\alpha E - \beta A)^T K$ and $K(\alpha A - \beta E)$ are nonsingular, and (3.5) holds. ■

As a consequence of the identity (3.6), if we select $s = \alpha$ and $t = \beta$, with α and β real and such that $\beta E - \alpha A$ is invertible, we find that the matrix

$$S = (\alpha E - \beta A)(\beta E - \alpha A)^{-1} \tag{3.7}$$

is symplectic, i.e., $S^T K S = K$. In other words, if the pair $(sE - tA, K)$ is symplectic, so $(sI_{2n} - tS, K)$. It is easy to verify that under the transformation (3.7), if z is an eigenvalue of $sE - tA$, then

$$z' = f(z) := \frac{\alpha - \beta z}{\beta - \alpha z} \tag{3.8}$$

is an eigenvalue of S. The transformation $f(\cdot)$ is an involution, since we have $f(f(z)) = z$, and it maps the unit circle into the unit circle. In particular, it maps $z = \pm 1$ into $z' = \mp 1$. The transformation (3.7) can be used to relate the analysis of symplectic pencils to the simpler case of symplectic matrices. To connect more precisely the eigenstructures of the pencils $sE - tA$ and $sI_{2n} - tS$, we employ the following lemma.

LEMMA 3.2. *Given a Jordan block $J_r(z)$ corresponding to $z \in \mathbb{C}$, if $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ are such that $z \neq \gamma/\delta$, the matrix*

$$J' = (\gamma I_r - \delta J_r(z))^{-1}(\alpha I_r - \beta J_r(z)) \tag{3.9}$$

is similar to $J_r(h(z))$, with

$$h(z) = \frac{\alpha - \beta z}{\gamma - \delta z}, \tag{3.10}$$

i.e., there exists an invertible similarity transformation T such that $J' = T^{-1}J_r(h(z))T$.

Proof. Let $D = \alpha\delta - \beta\gamma$ denote the determinant of the Moebius transformation (3.10). The matrix J' is upper triangular and can be expressed as

$$J' = h(z)I_r + \sum_{k=1}^{r-1} c_k (Z_r)^k \tag{3.11a}$$

with

$$c_k = D\frac{\delta^{k-1}}{(\gamma - \delta z)^{k+1}}. \tag{3.11b}$$

This shows that $h(z)$ is the only eigenvalue, with multiplicity r, of J'. Furthermore, the only right eigenvector of J' corresponding to $h(z)$ is $v = [1 \ 0 \ \ldots \ 0]^T$. This implies that J' admits a single Jordan block of size r corresponding to $h(z)$. ∎

Consider now the symplectic matrix S obtained by applying the transformation (3.7) to a symplectic pencil $sE - tA$. Let P^T be a $r \times 2n$ matrix whose rows span a left invariant subspace of S, so that

$$P^T S = J' P^T \tag{3.12}$$

where J' is an $r \times r$ matrix. The expression (3.7) implies

$$(\alpha I_r - \beta J')P^T E = (\beta I_r - \alpha J')P^T A, \tag{3.13}$$

so that the rows of P^T span a left *deflating subspace* (see [15] for the definition and properties of deflating subspaces) of the pencil $sE - tA$. This implies the following result.

LEMMA 3.3. *Let $sE - tA$ be a regular pencil. If, in the transformation (3.7), α and β are selected in such a way that $\beta E - \alpha A$ is invertible, then to each elementary divisor of the form $(s - z_0' t)^r$ with $z_0' \neq \beta/\alpha$ of $sI_n - tS$ corresponds an elementary divisor $(s - z_0 t)^r$ of $sE - tA$ with $z_0 = f(z_0')$, where $f(\cdot)$ is given by (3.8).*

Proof. In (3.9), let $J' = J_r(z_0')$ be a Jordan block of size r corresponding to the eigenvalue z_0'. Then (3.13) can be rewritten as $JP^TE = P^TA$, with

$$J = (\beta I_r - \alpha J_r(z_0'))^{-1}(\alpha I_r - \beta J_r(z_0')). \qquad (3.14)$$

But according to Lemma 3.2, J is similar to $J_r(z_0)$ with $z_0 = f(z_0')$. ■

Lemmas 3.1 and 3.3 can be used to give the following characterization of the elementary-divisor structure of symplectic pencils.

THEOREM 3.2. *If $(s_0 s - t_0 t)^r$ is an elementary divisor of a symplectic pencil $sE - tA$, so is $(t_0 s - s_0 t)^r$. Furthermore, the elementary divisors $(s \pm t)^{2r+1}$ of odd degrees corresponding to eigenvalues at ± 1 occur in pairs.*

Proof. The first part of the statement is a consequence of the equivalence of $sE - tA$ and $tE - sA$ (see also [20]). The second part uses the fact that since the matrix S obtained by (3.7) is symplectic, its elementary divisors of odd degree corresponding to eigenvalues at ± 1 must occur in pairs, as shown in [13]. Then, according to Lemma 3.3, the elementary divisors of $sE - tA$ corresponding to eigenvalues at ± 1 have the same property. ■

The results described below will also be useful in the sequel.

LEMMA 3.4. *The following equivalence relations hold:*

(1) *For any $r \in \mathbb{N}$ and any $a \in \mathbb{C} \setminus \{0\}$,*

$$sJ_r^T(a) - tI_r \sim sI_r - tJ_r(a^{-1}). \qquad (3.15)$$

(2) *For any $r \in \mathbb{N}$ and any $a \in \mathbb{C} \setminus \mathbb{R}$,*

$$sJ_{2r}^T(a, a^*) - tI_{2r} \sim sI_{2r} - tJ_{2r}(a^{-1}, a^{-*}). \qquad (3.16)$$

(3) *Let $r \in \mathbb{N}$, $\theta \in (0, \pi)$, $b = \tan(\theta/2)$, $E = I_{2r} - J_{2r}(jb, -jb)$, and $A = I_{2r} + J_{2r}(jb, -jb)$; then*

$$sE - tA \sim sI_{2r} - tJ_{2r}(e^{j\theta}, e^{-j\theta}). \qquad (3.17)$$

(4) *Let* $r \in \mathbb{N}$, $\epsilon = \pm 1$, $E = I_r - Z_r$, *and* $A = \epsilon(I_r + Z_r)$; *then*

$$sE - tA \sim sI_r - tJ_r(\epsilon). \tag{3.18}$$

Proof. (1): Clearly the equivalence relation (3.15) holds if and only if the matrix $J_r^{-T}(a)$ is similar to $J_r(a^{-1})$. Since any matrix is similar to its transpose, we only need to prove that $J_r^{-1}(a)$ is similar to $J_r(a^{-1})$. This is a consequence of Lemma 3.2, with $\alpha = 1$, $\beta = 0$, $\gamma = 0$, and $\delta = -1$.

(2): We need to show that the matrix $J_{2r}^{-T}(a, a^*)$ is similar to $J_{2r}(a^{-1}, a^{-*})$. But since the matrix $J_{2r}(a, a^*)$ is similar to $J_r(a) \oplus J_r(a^*)$ [9, p. 152], the proof follows from 1.

(3): We need to show that the matrix AE^{-1} is similar to $J_{2r}(e^{j\theta}, e^{-j\theta})$. We have

$$AE^{-1} = -2(-E)^{-1} - I, \tag{3.19}$$

as can be checked by right multiplication of the identity $A = 2I - E$ by E^{-1}. The right-hand side of (3.13) can be rewritten as $-2[J_{2r}(-1 + jb, -1 - jb)]^{-1} - I$. Invoking the same argument as for the derivation of item 2, the matrix $-2[J_{2r}(-1 + jb, -1 - jb)]^{-1}$ is similar to $J_{2r}(-2(-1 + jb)^{-1}, -2(-1 - jb)^{-1}) = J_{2r}((2 + j2b)(1 + b^2)^{-1}, (2 - j2b)(1 + b^2)^{-1})$, and hence AE^{-1} is similar to $J_{2r}((2 + j2b)(1 + b^2)^{-1} - 1, (2 - j2b)(1 + b^2)^{-1} - 1)$. This last expression, taking into account the definition of b, may be written as $J_{2r}(e^{j\theta}, e^{-j\theta})$.

(4): To prove that the matrix $\epsilon(I_r + Z_r)(I_r - Z_r)^{-1}$ is similar to $J_r(\epsilon)$, we apply Lemma 3.2 with $\alpha = -\beta = \epsilon$, $\gamma = \delta = 1$, and $z = 0$. This concludes the proof. ∎

LEMMA 3.5. *Let* $sE - tA$ *and* $sE' - tA'$ *be two regular matrix pencils, possibly of different sizes. Then the Lyapunov-type matrix equation*

$$(E')^T ME - (A')^T MA = 0 \tag{3.20}$$

admits a nonzero solution if and only if there exists two eigenvalues $z_0 = s_0/t_0$ *and* $z_0' = s_0'/t_0'$ *of* $sE - tA$ *and* $sE' - tA'$, *respectively, such that* $s_0' s_0 = t_0' t_0$, *or equivalently*

$$z_0 = (z_0')^{-1}. \tag{3.21}$$

For a proof of this lemma, we refer to [15].

4. PROOF OF THE MAIN RESULT

4.1. *Introduction*

In this section we prove Theorem 2.1. In particular, we shall prove that there exist two real invertible matrices V and W such that

$$(sE - tA)V = W\tilde{P}(s,t) \tag{4.1}$$

and

$$W^T K W - \tilde{K}, \tag{4.2}$$

where $\tilde{P}(s,t)$ and $\tilde{K}$ have the form specified by Theorem 2.1. To complete the proof, it will be shown that if a symplectic pair $(P(s,t), K)$ is symplectically equivalent to $(\tilde{P}_1(s,t), \tilde{K}_1)$ and to $(\tilde{P}_2(s,t), \tilde{K}_2)$, where both $(\tilde{P}_1(s,t), \tilde{K}_1)$ and $(\tilde{P}_2(s,t), \tilde{K}_2)$ have the form specified by Theorem 2.1, then

$$\left(\tilde{P}_1(s,t), \tilde{K}_1\right) = \left(\tilde{P}_2(s,t), \tilde{K}_2\right). \tag{4.3}$$

The approach employed to derive Theorem 2.1 is constructive and allows, in principle, the computation of the matrices V and W. An alternative approach based on Theorem 2.1 of [20] and Theorem 3.5.2 of [12] would consist first of showing that $(P(s,t), K)$ is symplectically equivalent to a pair of the form $(P_0(s,t), K_0)$, with

$$P_0(s,t) = s\begin{bmatrix} I & 0 & 0 \\ 0 & I_m & 0 \\ 0 & 0 & N \end{bmatrix} - t\begin{bmatrix} S & 0 & 0 \\ 0 & N^T & 0 \\ 0 & 0 & I_m \end{bmatrix}, \tag{4.4a}$$

$$K_0 = \begin{bmatrix} K_1 & 0 & 0 \\ 0 & 0 & -I_m \\ 0 & I_m & 0 \end{bmatrix}, \tag{4.4b}$$

where m denotes the algebraic multiplicity of the zero eigenvalues of $P(s,t)$, N is a real nilpotent matrix of size $m \times m$, K_1 is a real skew-symmetric matrix of size $2(n-m)$, and S is K_1-symplectic, i.e., $S^T K_1 S = K_1$. Then, one could adapt the results of Cikunov [5, 4] and Laub and Meyer [13] to bring the pair $(sI - tS, K_1)$ to the desired canonical form. However, the blocks of types 1–4 considered by Cikunov and by Laub and Meyer do not have the same structure as those considered here, so that additional transformations would be needed to bring them to the desired form. As a consequence, the computations required by this alternative derivation do not appear significantly shorter than those appearing below.

As a first step, observe that if $K = -K^T$ is a real nonsingular skew-symmetric matrix of even size, there exists a nonsingular matrix W_0 such that $W_0^T K W_0$ has the structure (1.1). Then, without loss of generality, we may assume that, in the symplectic pair $(P(s,t), K)$, the matrix K is given by (1.1).

Consider now the Weierstrass canonical decomposition of the symplectic pencil $sE - tA$. According to Theorem 3.2, each block of the decomposition corresponding to an eigenvalue at $a \in \mathbb{C}$ with $a \neq \pm 1$ is matched by an identical block at a^{-1}. For the particular case of an eigenvalue at $a_0 = 0$ whose block $s\mathscr{E}_0 - t\mathscr{A}_0$ takes the form

$$\mathscr{E}_0 = \bigoplus_{k=1}^{p_0} I_{r_k}, \qquad \mathscr{A}_0 = \bigoplus_{k=1}^{p_0} J_{r_k}(0), \tag{4.5}$$

the matching block is the block $s\mathscr{E}_\infty - t\mathscr{A}_\infty$ corresponding to the eigenvalue at $a_0^{-1} = \infty$, with the structure

$$\mathscr{E}_\infty = \bigoplus_{k=1}^{p_0} J_{r_k}(0), \qquad \mathscr{A}_\infty = \bigoplus_{j=1}^{p_0} I_{r_k}, \tag{4.6}$$

where the integers r_k in (4.5) and (4.6) are the same.

We can regroup the blocks of the Weierstrass canonical form corresponding to (1) reciprocal real eigenvalues a, a^{-1} (2) complex quadruples a, a^*, a^{-1}, a^{-*}, or (3) complex conjugate pairs $a = e^{j\theta}$, $a^* = e^{-j\theta}$ on the unit circle, where for all three such cases we assume $a \neq \pm 1$. By using the equivalence relations of Lemma 3.4, it is easy to verify that under equivalence, each grouping can be brought to the form specified for blocks of types 1, 2, and 3 in the canonical form of Theorem 2.1. Next, consider the blocks of the Weierstrass canonical form corresponding to eigenvalues at $a = \pm 1$. According to Theorem 3.2, the Jordan blocks of odd size corresponding to such eigenvalues appear in pairs, so that under equivalence they can be brought to the form specified for blocks of type 4.

At this point, we have show that there exist real matrices V and W such that the equivalence relation (4.1) holds. We now demonstrate that V and W may be chosen in such a way that the identity (4.2) is satisfied. First, we prove that $G := W^T K W$ is block-diagonal. To do so, let

$$E_D := \bigoplus_{i=1}^{l} E_i, \qquad A_D := \bigoplus_{i=1}^{l} A_i. \tag{4.7}$$

By using the symplectic property (1.5) of the pencil $sE - tA$, the equivalence relation (4.1) implies

$$E_D^T G E_D = A_D^T G A_D. \tag{4.8}$$

Partitioning G as $G = [G_{ij}]$, where G_{ij} has the same row size as E_i and the same column size as E_j, the equation (4.8) can be rewritten block by block as

$$E_i^T G_{ij} E_j - A_i^T G_{ij} A_j = 0. \tag{4.9}$$

But, for $i \neq j$, the eigenvalues of $sE_i - tA_i$ and $sE_j - tA_j$ never satisfy the condition (3.21), so that, in view of Lemma 3.5, $G_{ij} = 0$ for $i \neq j$. As a consequence, G admits the block-diagonal structure

$$G := W^T K W = \bigoplus_{i=1}^{l} G_i, \tag{4.10}$$

where the blocks G_i satisfy

$$E_i^T G_i E_i - A_i^T G_i A_i = 0. \tag{4.11}$$

We now show that the matrices W and V can be selected so that $G_i = K_i$, where, depending on whether we are considering blocks of type 1, 2, 3, or 4, K_i admits the structure (2.2c), (2.3c), (2.4c), or (2.5c). The key step consists in showing that the solution G_i of (4.11) can be expressed as

$$G_i = Q_i^T K_i Q_i, \tag{4.12}$$

where K_i takes the form (2.2c), (2.3c), (2.4c), or (2.5c), depending on the block type, and where Q_i commutes with the matrices E_i and A_i. As a consequence, if we define

$$\overline{V}_i := Q_i^{-1}, \qquad \overline{W}_i := Q_i^{-1}, \tag{4.13}$$

the relation

$$(sE_i - tA_i)\bar{V}_i = \bar{W}_i(sE_i - tA_i) \tag{4.14}$$

holds.

4.2. Justification of (4.12)

Type 1. Let (a_i, a_i^{-1}) with $|a_i| < 1$ be a pair of real eigenvalues corresponding to a block $sE_i - tA_i$ with

$$E_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} I_{r_k} & 0 \\ 0 & J_{r_k}^T \end{bmatrix}, \tag{4.15a}$$

$$A_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} J_{r_k} & 0 \\ 0 & I_{r_k} \end{bmatrix}, \tag{4.15b}$$

where for simplicity we denote $J_{r_k} = J_{r_k}(a_i)$. Partitioning G_i as $G_i = [G_{kl}]$, the (k, l)th block of Equation (4.11) can be written as

$$\begin{bmatrix} I_{r_k} & 0 \\ 0 & J_{r_k} \end{bmatrix} G_{kl} \begin{bmatrix} I_{r_l} & 0 \\ 0 & J_{r_l}^T \end{bmatrix} = \begin{bmatrix} J_{r_k}^T & 0 \\ 0 & I_{r_k} \end{bmatrix} G_{kl} \begin{bmatrix} J_{r_l} & 0 \\ 0 & I_{r_l} \end{bmatrix}. \tag{4.16}$$

Taking into account Lemma 3.5, this equation implies that G_{kl} has the structure

$$G_{kl} = \begin{bmatrix} 0 & G_{kl}^1 \\ G_{kl}^2 & 0 \end{bmatrix}, \tag{4.17}$$

where

$$G_{kl}^1 J_{r_l}^T = J_{r_k}^T G_{kl}^1, \tag{4.18a}$$

$$J_{r_k} G_{kl}^2 = G_{kl}^2 J_{r_l}. \tag{4.18b}$$

Since G_{kl} is skew-symmetric, we have

$$G_{kl} = \begin{bmatrix} 0 & G_{kl}^{1} \\ G_{kl}^{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & -G_{lk}^{2T} \\ -G_{lk}^{1T} & 0 \end{bmatrix} = -G_{lk}^{T}. \tag{4.19}$$

In particular, for $k = l$,

$$G_{kk}^{1} = -G_{kk}^{2T}. \tag{4.20}$$

Now define the matrix $Q_i = [Q_{kl}]$ partitioned conformally with $G_i[G_{kl}]$, with

$$Q_{kl} := \begin{bmatrix} G_{kl}^{2} & 0 \\ 0 & \overline{Q}_{kl} \end{bmatrix}, \tag{4.21a}$$

where the block $\overline{Q}_{kl}$ is given by

$$\overline{Q}_{kl} = \begin{cases} 0_{r_k \times r_l} & \text{for} \quad k \neq 1, \\ I_{r_k} & \text{for} \quad k = l, \end{cases} \tag{4.21b}$$

and where $0_{r_k \times r_l}$ denotes the zero matrix of size $r_k \times r_l$.

It is not difficult to check that (4.12) holds, where K_i is given by (2.2c). Specifically,

$$[K_i Q_i]_{kl} = K_{kk} Q_{kl} = \begin{bmatrix} 0 & -\overline{Q}_{kl} \\ G_{kl}^{2} & 0 \end{bmatrix}, \tag{4.22a}$$

$$[Q_i^T]_{kl} = Q_{lk}^T = \begin{bmatrix} G_{lk}^{2T} & 0 \\ 0 & \overline{Q}_{lk}^T \end{bmatrix}, \tag{4.22b}$$

and

$$[Q_i^T K_i Q_i]_{kl} = \sum_{h=1}^{p_i} [Q_i^T]_{kh} [K_i Q_i]_{hl} = [Q_k^T]_{kk} [K_i Q_i]_{kl} + [Q_i^T]_{kl} [K_i Q_i]_{ll}$$

$$= \begin{bmatrix} 0 & -G_{lk}^{2T} \\ G_{kl}^{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & G_{kl}^{1} \\ G_{kl}^{2} & 0 \end{bmatrix}. \tag{4.23}$$

Furthermore, Q_i commutes with E_i and A_i. This can be verified by noting from (4.18b) that

$$[Q_i A_i]_{kl} = Q_{kl} A_{ll} = \begin{bmatrix} G_{kl}^2 J_{r_l} & 0 \\ 0 & \overline{Q}_{kl} \end{bmatrix} = \begin{bmatrix} J_{r_k} G_{kl}^2 & 0 \\ 0 & \overline{Q}_{kl} \end{bmatrix} = [A_i Q_i]_{kl}. \tag{4.24}$$

Similarly, we can check that Q_i commutes with E_i.

Type 2. Let $(a_i, a_i^*, a_i^{-1}, a_i^{-*})$ with $a_i = \sigma_i + j\omega_i$ and $|a_i| < 1$ be a quadruple of complex eigenvalues corresponding to the block $sE_i - tA_i$ with

$$E_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} I_{2r_k} & 0 \\ 0 & J_{2r_k}^T \end{bmatrix}, \tag{4.25a}$$

$$A_i = \bigoplus_{k=1}^{p_i} \begin{bmatrix} J_{2r_k} & 0 \\ 0 & I_{2r_k} \end{bmatrix}, \tag{4.25b}$$

where J_{2r_k} is an abbreviation for $J_{2r_k}(a_i, a_i^*)$. Let also G_i be the corresponding diagonal block of G. By employing the same argument as for type 1 blocks, it can be shown that there exists a matrix Q_i commuting with E_i and A_i such that (4.12) holds, where K_i has the structure (2.3c).

Type 3. Let $e^{\pm j\theta_i}$ be a pair of complex conjugate eigenvalues on the unit circle corresponding to the block $sE_i - tA_i$ with

$$E_i = \bigoplus_{k=1}^{p_i} \left(I_k - J_{2r_k}\right), \tag{4.26a}$$

$$A_i = \bigoplus_{k=1}^{p_i} \left(I_k + J_{2r_k}\right), \tag{4.26b}$$

where J_{2r_k} is an abbreviation for $J_{2r_k}(jb_i, -jb_i)$ with $b_i := \tan(\theta_i/2)$. Partitioning G_i as $G_i = [G_{kl}]$, the equation (4.11) can be rewritten in block form as

$$J_{2r_k}^T G_{kl} + G_{kl} J_{2r_l} = 0. \tag{4.27}$$

By observing that

$$J_{2r}(jb, -jb) = bI_r \otimes \Sigma_2 + Z_r \otimes I_2, \tag{4.28a}$$

we deduce that G_{kl} can be represented as

$$G_{kl} = \Gamma_{kl}^R \otimes (\Sigma_2)^{r_{k,l}} + \Gamma_{kl}^I \otimes (\Sigma_2)^{r_{k,l}+1} \tag{4.28b}$$

with $r_{k,l} = \max(r_k, r_l)$, where each of the $r_k \times r_l$ matrices Γ_{kl}^R and Γ_{kl}^I satisfies an equation of the form

$$Z_{r_k}^T \Gamma_{kl} + \Gamma_{kl} Z_{r_l} = 0. \tag{4.29}$$

Noting that $Z_r^T = -\Sigma_r Z_r \Sigma_r^T$, the equation (4.29) can be rewritten as

$$Z_{r_k} \Sigma_{r_k}^T \Gamma_{kl} = \Sigma_{r_k}^T \Gamma_{kl} Z_{r_l}, \tag{4.30}$$

from which we conclude that if Γ_{kl} solves (4.29), it can be expressed as

$$\Gamma_{kl} = \Sigma_{r_k} p_{kl}(Z_{r_k}) \begin{bmatrix} 0_{(r_k - r_l) \times r_l} \\ I_{r_l} \end{bmatrix} \tag{4.31a}$$

for $r_k \geqslant r_l$, where $p_{kl}(x)$ is a polynomial of degree less or equal to $r_k - 1$ which is divisible by $x^{r_k - r_l}$. Similarly, for $r_l \geqslant r_k$ we have

$$\Gamma_{kl} = \begin{bmatrix} 0_{r_k \times (r_l - r_k)} & I_{rk} \end{bmatrix} \Sigma_{r_l} P_{kl}(Z_{r_l}) \tag{4.31b}$$

where $p_{kl}(x)$ is a polynomial of degree less or equal to $r_l - 1$ which is divisible by $x^{r_l - r_k}$.

Since G_i is skew-symmetric, its blocks must satisfy $G_{kl} = -G_{lk}^T$. Taking into account the identity

$$(Z_r^T)^k \Sigma_r = (-1)^k \Sigma_r (Z_r)^k, \tag{4.32}$$

this implies that the polynomials $p_{kl}^R(x)$ and $p_{kl}^I(x)$ parametrizing Γ_{kl}^R and Γ_{kl}^I must satisfy

$$p_{kl}^R(x) = p_{lk}^R(-x), \qquad p_{kl}^I(x) = -p_{lk}^I(-x). \tag{4.33}$$

In particular, for $k = l$, we have

$$p_{kk}^R(x) = p_{kk}^R(-x), \qquad p_{kk}^I(x) = -p_{kk}^I(-x), \tag{4.34}$$

so that $p_{kk}^R(x)$ and $p_{kk}^I(x)$ are respectively even and odd polynomials of degree less than or equal to $r_k - 1$.

Now that we have completely characterized the structure of G_i, all we need to do is construct a matrix Q_i such that the factorization (4.12) holds, where K_i has the structure (2.4c) and Q_i commutes with both E_i and A_i. Without loss of generality, we can assume that the blocks appearing in the decomposition (4.26a)–(4.26b) have decreasing sizes, so that $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_{p_i}$. Partitioning $Q_i = [Q_{kl}]$ in accordance with this decomposition, we require Q_i to be *block upper triangular*, i.e., $Q_{kl} = 0$ for $k > l$, and

$$Q_{kl} = Q_{kl}^R \otimes (\Sigma_2)^{r_k} + Q_{kl}^I \otimes (\Sigma_2)^{r_k + 1} \tag{4.35}$$

for $k \leqslant l$, where the matrices Q_{kl}^R and Q_{kl}^I have the structure

$$Q_{kl}^R = u_{kl}^R(Z_{r_k}) \begin{bmatrix} 0_{(r_k - r_l) \times r_l} \\ I_{r_l} \end{bmatrix} \tag{4.36a}$$

$$Q_{kl}^I = u_{kl}^I(Z_{r_k}) \begin{bmatrix} 0_{(r_k - r_l) \times r_l} \\ I_{r_l} \end{bmatrix}. \tag{4.36b}$$

In this representation, $u_{kl}^R(x)$ and $u_{kl}^I(x)$ are polynomials of degree less or equal to $r_k - 1$ which are required to be divisible by $x^{r_k - r_l}$. For $k = l$, we also require that the polynomials $u_{kk}^R(x)$ and $u_{kk}^I(x)$ should be even and odd, respectively.

To evaluate the polynomials $u_{kl}^R(x)$ and $u_{kl}^I(x)$ so that the factorization (4.12) holds, we need only to compare the representation of each block on both sides of (4.12). To do so, it is convenient to introduce the complex polynomials

$$p_{kl}(x) = p_{kl}^R(x) + jp_{kl}^I(x), \tag{4.37a}$$

$$u_{kl}(x) = u_{kl}^R(x) + ju_{kl}^I(x), \tag{4.37b}$$

as well as the $p_i \times p_i$ rational matrix

$$\Phi_i(x) = \left[\frac{p_{kl}(x)}{(jx)^{r_{k,l}}} \right] \tag{4.38}$$

and the $p_i \times p_i$ polynomial matrix $U_i(x) = [u_{kl}(x)]$. Then, the property (4.33) of the polynomials $p_{kl}^R(x)$ and $p_{kl}^I(x)$ implies that $\Phi_i(x)$ has the *para-Hermitian property*

$$(\Phi_i(-x^*))^H = \Phi_i(x), \tag{4.39}$$

where H denotes the Hermitian transpose. Furthermore, because of the upper triangular structure of Q_i, the polynomial matrix $U_i(x)$ is *upper triangular*. Note also that the requirement that $u_{kk}^R(x)$ and $u_{kk}^I(x)$ should be even and odd respectively is equivalent to requiring that the polynomials $u_{kk}(x)$ appearing on the diagonal of $U_i(x)$ should have the para-Hermitian property

$$(u_{kk}(-x^*))^* = u_{kk}(x). \tag{4.40}$$

Let $S_i(x)$ be the rational para-Hermitian matrix

$$S_i(x) = \operatorname{diag}\left\{ \frac{\kappa_k}{(jx)^{r_k}}; 1 \leqslant k \leqslant p_i \right\}. \tag{4.41}$$

Then, by identifying the (k, l)th block on each side of the factorization (4.12), it is easy to verify that it is equivalent to the partial rational matrix factorization

$$\Phi_i(x) = \pi_-\left\{ (U_i(-x^*))^H S_i(x) U_i(x) \right\}, \tag{4.42}$$

where the operator $\pi_-\{\cdot\}$ projects a rational matrix function onto its strictly proper part.

The existence of such a factorization is a consequence of the para-Hermitian property of $\Phi_i(x)$. In this context, it is useful to note that because the matrices W and K are invertible, the blocks G_i in (4.10) are invertible. It is shown in Appendix B that the invertibility of G_i implies that when $\Phi_i(x)$ is express as

$$\Phi_i(x) = D_i^{-1}(x)N_i(x) = \left(N_i(-x^*)\right)^H D_i^{-1}(x) \tag{4.43a}$$

with

$$D_i(x) = \operatorname{diag}\left\{(jx)^{r_k};\ 1 \leqslant k \leqslant p_i\right\}, \tag{4.43b}$$

the constant matrix $N_i(0)$ is invertible, so that (4.43a) yields left- and right-coprime matrix fraction descriptions of $\Phi_i(x)$. The reader is referred to [11] for a study of the properties of matrix function descriptions of a rational matrix. But (4.42) indicates that $(U_i(-x^*))^H S_i(x) U_i(x)$ is, up to the addition of a polynomial matrix, a polynomial matrix description of $\Phi_i(x)$ of the same order as the irreducible representations (4.43a). It is therefore also irreducible, so that the constant matrix $U_i(0)$ is invertible, i.e., the constant term of the polynomial $u_{kk}(x)$ is nonzero for all k. Then, the factorization (4.42) can be performed row by row, in a manner similar to the LDU factorization of a symmetric matrix. Thus, for $k \leqslant l$, let

$$\tilde{p}_{kl}(x) := p_{kl}(x) - \sum_{h=1}^{k-1} (jx)^{r_k - r_h} (u_{hk}(-x^*))^* u_{hl}(x). \tag{4.44}$$

Note that $\tilde{p}_{kk}(x)$ is para-Hermitian. Also, if $p_{kl}(x)$ is divisible by $x^{r_k - r_l}$ and $u_{hl}(x)$ is divisible by $x^{r_h - r_l}$ for $h < l$, then $\tilde{p}_{kl}(x)$ is also divisible by $x^{r_k - r_l}$. Then, for $1 \leqslant k \leqslant p_i$ and $l \leqslant l \leqslant p_i$, the factorization (4.42) is obtained by solving the equations

$$\kappa_k u_{kk}^2(x) = \tilde{p}_{kk}(x) \bmod x^{r_k} \tag{4.45a}$$

$$\kappa_k u_{kk}(x) u_{kl}(x) = \tilde{p}_{kl}(x) \bmod x^{r_k}, \tag{4.45b}$$

where we have used the para-Hermitian property of $u_{kk}(x)$, and where the modulo operation indicates that the polynomials on both sides of the above identities need only to be equal modulo the addition of a multiple of x^{r_k}. From (4.45a), we see that κ_k is just the sign of the constant term of $\tilde{p}_{kk}(x)$. Then by matching the coefficients of increasing orders on both sides of (4.45a), it is easy to check this equation admits a unique solution. Substituting the polynomial $u_{kk}(x)$ inside (4.45b), we can then determine the successive coefficients of $u_{kl}(x)$ for all $l > k$. Note also that since the constant coefficient of $u_{kk}(x)$ is nonzero, and $\tilde{p}_{kl}(x)$ is divisible by $x^{r_k - r_l}$, $u_{kl}(x)$ will also be divisible by $x^{r_k - r_l}$.

This yields the factorization (4.12). Finally, from the structure of the blocks Q_{kl}, it is clear that Q_i commutes with $\oplus_{k=1}^{p_i} J_{2r_k}$, and thus with E_i and A_i.

Type 4. Let $\epsilon_i = \pm 1$ be an eigenvalue of the pencil corresponding to a block $sE_i - tA_i$ with

$$E_i = E_e \oplus E_o, \tag{4.46a}$$

$$A_i = A_e \oplus A_o, \tag{4.46b}$$

where the blocks of even size take the form

$$E_e := \bigoplus_{k=1}^{p_{ei}} \left(I_{2r_k} - Z_{2r_k} \right), \tag{4.47a}$$

$$A_e := \epsilon_i \bigoplus_{k=1}^{p_{ei}} \left(I_{2r_k} + Z_{2r_k} \right), \tag{4.47b}$$

and the blocks of odd size are given by

$$E_0 := \bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} I_{2r'_k+1} - Z_{2r'_k+1} & 0 \\ 0 & \left(I_{2r'_k+1} + Z_{2r'_k+1} \right)^T \end{bmatrix}, \tag{4.48a}$$

$$A_o := \epsilon_i \bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} I_{2r'_k+1} + Z_{2r'_k+1} & 0 \\ 0 & \left(I_{2r'_k+1} - Z_{2r'_k+1} \right)^T \end{bmatrix}. \tag{4.48b}$$

Then, we can partition G_i as

$$G_i := \begin{bmatrix} G_e & G_{eo} \\ -G_{eo}^T & G_o \end{bmatrix}, \tag{4.49}$$

where the sizes of the even and odd blocks match those of the corresponding blocks of E_i and A_i in (4.46a) and (4.46b). The equation (4.11) reduces to the three block equations

$$E_e^T G_e E_e = A_e^T G_e A_e, \tag{4.50a}$$

$$E_e^T G_{eo} E_o = A_e^T G_{eo} A_o, \tag{4.50b}$$

$$E_o^T G_o = A_o^T G_o A_o. \tag{4.50c}$$

Each of these equations can be written in a blockwise manner by subdividing G_e, G_{eo} and G_o as $G_e = [G_{kl}^e]$, $G_{eo} = [G_{kl}^{eo}]$, and $G_o = [G_{kl}^o]$. The equations for each block are then given by

$$Z_{2r_k}^T G_{kl}^e + G_{kl}^e Z_{2r_k} = 0, \tag{4.51a}$$

$$Z_{2r_k}^T G_{kl}^{eo} + G_{kl}^{eo} \begin{bmatrix} Z_{2r_l'+1} & 0 \\ 0 & -Z_{2r_l'+1}^T \end{bmatrix} = 0, \tag{4.51b}$$

$$\begin{bmatrix} Z_{2r_k'+1}^T & 0 \\ 0 & -Z_{2r_k'+1} \end{bmatrix} G_{kl}^o + G_{kl}^o \begin{bmatrix} Z_{2r_l'+1} & 0 \\ 0 & -Z_{2r_l'+1}^T \end{bmatrix} = 0. \tag{4.51c}$$

From these identities, we deduce that the blocks composing G_e, G_{eo}, and G_o have the following structure.

Even blocks. The blocks G_{kl}^e admit the parametrization

$$G_{kl}^e = \Sigma_{2r_k} p_{kl}^3 \left(Z_{2r_k} \right) \begin{bmatrix} 0_{2(r_k - r_l) \times 2r_l} \\ I_{2r_l} \end{bmatrix} \tag{4.52a}$$

for $r_k \geqslant r_l$, and

$$G_{kl}^e = \begin{bmatrix} 0_{2r_k \times 2(r_l - r_k)} & I_{2r_k} \end{bmatrix} \Sigma_{2r_l} p_{kl}^e \left(Z_{2r_l} \right) \tag{4.52b}$$

for $r_l \geqslant r_k$ where $p_{kl}^e(x)$ is a polynomial of degree less than or equal to $2\max(r_k, r_l) - 1$ which is divisible by $x^{2|r_k - r_l|}$. The skew symmetry of G_e implies also

$$G_{kl}^e = -G_{lk}^{eT}, \tag{4.53}$$

from which we conclude that the polynomials $p_{kl}^{e}(x)$ satisfy

$$p_{kl}^{e}(x) = p_{lk}^{e}(-x). \tag{4.54}$$

In particular, for $k = l$ we have

$$p_{kk}^{e}(x) = p_{kk}^{e}(-x), \tag{4.55}$$

so that $p_{kk}^{e}(x)$ is an even polynomial of degree less or equal to $2(r_k - 1)$.

Cross blocks. By partitioning G_{kl}^{eo} as

$$G_{kl}^{eo} = \begin{bmatrix} G_{kl}^{eo1} & G_{kl}^{eo2} \end{bmatrix}, \tag{4.56}$$

we find that

$$G_{kl}^{eo1} = \Sigma_{2r_k} p_{kl}^{eo1}(Z_{2r_k}) \begin{bmatrix} 0_{[2r_k-(2r_l'+1)]\times(2r_l'+1)} \\ I_{2r_l'+1} \end{bmatrix}, \tag{4.57a}$$

$$G_{kl}^{eo2} = p_{kl}^{eo2}(-Z_{2r_k}^{T}) \begin{bmatrix} I_{2r_l'+1} \\ 0_{[2r_k-(2r_l'+1)]\times(2r_l'+1)} \end{bmatrix} \tag{4.57b}$$

for $2r_k \geqslant 2r_l' + 1$, and

$$G_{kl}^{eo1} = \begin{bmatrix} 0_{2r_k\times((2r_l'+1)-2r_k)} & I_{2r_k} \end{bmatrix} \Sigma_{2r_l'+1} p_{kl}^{eol}(Z_{2r_l'-1}), \tag{4.58a}$$

$$G_{kl}^{eo2} = \begin{bmatrix} 0_{2r_k\times((2r_l'+1)-2r_k)} & I_{2r_k} \end{bmatrix} p_{kl}^{eo2}(-Z_{2r_l'+1}^{T}), \tag{4.58b}$$

for $2r_l' + 1 \geqslant 2r_k$, where for $j = 1, 2$, $p_{kl}^{eoj}(x)$ is a polynomial of degree less than or equal to $\max(2r_k - 1, 2r_l')$ which is divisible by $x^{|2r_k-(2r_l'+1)|}$.

Odd blocks. The blocks

$$G_{kl}^{o} = \begin{bmatrix} G_{kl}^{o1} & G_{kl}^{o2} \\ G_{kl}^{o3} & G_{kl}^{o4} \end{bmatrix} \tag{4.59}$$

have the structure

$$G_{kl}^{o1} = p_{kl}^{o1}\left(-Z_{2r_k'+1}^T\right)\Sigma_{2r_k'+1}\begin{bmatrix} 0_{2(r_k'-r_l')\times(2r_l'+1)} \\ I_{2r_l'+1} \end{bmatrix}, \tag{4.60a}$$

$$G_{kl}^{o2} = p_{kl}^{o2}\left(-Z_{2r_k'+1}^T\right)\begin{bmatrix} I_{2r_l'+1} \\ 0_{2(r_k'-r_l')\times(2r_l'+1)} \end{bmatrix}, \tag{4.60b}$$

$$G_{kl}^{o3} = P_{kl}^{o3}\left(Z_{2r_k'+1}\right)\begin{bmatrix} 0_{2(r_k'-r_l')\times 2r_l'+1)} \\ I_{2r_l'+1} \end{bmatrix}, \tag{4.60c}$$

$$G_{kl}^{o4} = p_{kl}^{o4}\left(Z_{2r_k'+1}\right)\Sigma_{2r_k'+1}\begin{bmatrix} I_{2r_l'+1} \\ 0_{2(r_k'-r_l')\times(2r_l'+1)} \end{bmatrix} \tag{4.60d}$$

for $r_k' \geqslant r_l'$, and

$$G_{kl}^{o1} = \begin{bmatrix} 0_{(2r_k'+1)\times 2(r_l'-r_k')} & I_{2r_k'+1} \end{bmatrix} p_{kl}^{o1}\left(-Z_{2r_l'+1}^T\right)\Sigma_{2r_l'+1}, \tag{4.61a}$$

$$G_{kl}^{o2} = \begin{bmatrix} 0_{(2r_k'+1)\times 2(r_l'-r_k')} & I_{2r_k'+1} \end{bmatrix} p_{kl}^{o2}\left(-Z_{2r_l'+1}^T\right), \tag{4.61b}$$

$$G_{kl}^{o3} = \begin{bmatrix} I_{2r_k'+1} & 0_{(2r_k'+1)\times 2(r_l'-r_k')} \end{bmatrix} p_{kl}^{o3}\left(Z_{2r_l'+1}\right), \tag{4.61c}$$

$$G_{kl}^{o4} = \begin{bmatrix} I_{2r_k'+1} & 0_{(2r_k'+1)\times 2(r_l'-r_k')} \end{bmatrix} p_{kl}^{o4}\left(Z_{2r_l'+1}\right)\Sigma_{2r_l'+1} \tag{4.61d}$$

for $r_l' \geqslant r_k'$. In these expressions, $p_{kl}^{oj}(x)$ with $j = 1, 2, 3, 4$ denotes a polynomial of degree less than or equal to $2\max(r_k', r_l')$ which is divisible by $x^{2|r_k'-r_l'|}$. The skew symmetry of the matrix G_o implies

$$G_{kl}^{o} = \begin{bmatrix} G_{kl}^{o1} & G_{kl}^{o2} \\ G_{kl}^{o3} & G_{kl}^{o4} \end{bmatrix} = -\begin{bmatrix} G_{lk}^{o1T} & G_{lk}^{o3T} \\ G_{lk}^{o2T} & G_{lk}^{o4T} \end{bmatrix} = -G_{lk}^{oT}, \tag{4.62}$$

from which we deduce that the polynomials $p_{kl}^{oj}(x)$ satisfy

$$p_{kl}^{o1}(x) = -p_{lk}^{o1}(-x), \tag{4.63a}$$

$$p_{kl}^{o2}(x) = -p_{lk}^{o3}(-x), \tag{4.63b}$$

$$p_{kl}^{o4}(x) = -p_{lk}^{o4}(-x). \tag{4.63c}$$

In particular, for $k = l$, this implies that for $j = 1, 4$

$$p_{kk}^{oj}(x) = -p_{kk}^{oj}(-x), \tag{4.64}$$

so that $p_{kk}^{o2}(x)$ and $p_{kk}^{o4}(x)$ are odd polynomials.

The construction of the factorization (4.12) for G_i can then be performed in two steps.

Step 1: We factor the even block G_e as

$$G_e = Q_e^T K_e Q_e, \tag{4.65a}$$

where

$$K_e := \bigoplus_{k=1}^{p_{ei}} \kappa_k \Sigma_{2r_k} \tag{4.65b}$$

is the even block of K_i in (2.5c). The existence of such a factorization relies on the invertibility of G_e, which is provided in Appendix A. Without loss of generality, we can assume that the blocks of G_e have decreasing size, with $2r_1 \geqslant 2r_2 \cdots \geqslant 2r_{p_{ei}}$. Then, partitioning Q_e as $Q_e = [Q_{kl}^e]$, we require Q_e to be block upper triangular with $Q_{kl}^e = 0$ for $k > l$ and

$$Q_{kl}^e = u_{kl}^e(Z_{2r_k}) \begin{bmatrix} 0_{2(r_k - r_l) \times 2r_l} \\ I_{2r_l} \end{bmatrix} \tag{4.66}$$

for $l \geqslant k$, where the polynomials $u_{kl}^e(x)$ are required to have degree less or equal to $2(r_k - 1)$ and to be divisible by $x^{2(r_k - r_l)}$. For $k = l$, we also require that the polynomials $u_{kk}^e(x)$ should be even.

If we introduce the $p_{ei} \times p_{ei}$ polynomial matrix $U_e(x) = [u_{kl}^e(x)]$, and the $p_{ei} \times p_{ei}$ rational matrix

$$\Phi_e(x) = \left[\frac{p_{kl}^e(x)}{x^{2r_{k,l}}} \right], \tag{4.67}$$

the property (4.54) of the polynomials $p_{kl}^e(x)$ implies that $\Phi_e(x)$ is *parasymmetric*, i.e.,

$$\Phi_e(x) = (\Phi_e(-x))^T, \tag{4.68}$$

where we assume that x is real. Because Q_e is block upper triangular, the polynomial matrix $U_e(x)$ is upper triangular. Then, substituting the parametrization (4.66) of the blocks Q_{kl}^e inside the factorization (4.65a), and introducing the rational parasymmetric matrix

$$S_e(x) = \operatorname{diag}\left\{ \frac{\kappa_k}{x^{2r_k}}; 1 \leqslant k \leqslant p_{ei} \right\}, \tag{4.69}$$

we can rewrite (4.65a) as the partial rational matrix factorization

$$\Phi_e(x) = \pi_- \left\{ (U_e(-x))^T S_e(x) U_e(x) \right\}. \tag{4.70}$$

This factorization is similar to the factorization (4.42) employed for blocks of type 3 and can be constructed in the same manner.

Note also that because G_e is invertible, Q_e is invertible, and the form (4.66) implies it commutes with E_e and A_e. Then, G_i admits the block LDU decomposition

$$G_i = \tilde{Q}_i^T \tilde{G}_i \tilde{Q}_i \tag{4.71}$$

with

$$\tilde{Q}_i := \begin{bmatrix} Q_e & K_e^{-1} Q_e^{-T} G_{eo} \\ 0 & I \end{bmatrix}, \tag{4.72a}$$

where

$$\tilde{G}_i = \begin{bmatrix} K_e & 0 \\ 0 & F_o \end{bmatrix} := \begin{bmatrix} K_e & 0 \\ 0 & G_o + G_{eo}^T G_e^{-1} G_{eo} \end{bmatrix} \tag{4.72b}$$

is skew-symmetric. An important feature of this expression (4.72a) for $\tilde{Q}_i$ is that because Q_e is block upper triangular, so is Q_e^{-1}. In addition, if we partition $Q_i^{-1} = [V_{kl}^e]$, it is easy to verify that the blocks V_{kl}^e admit a parametrization of the same type as the blocks Q_{kl}, except that the polynomials $u_{kl}^e(x)$ are replaced by the polynomials $v_{kl}^e(x)$. For $k \geqslant l$, the polynomial $v_{kl}^e(x)$ is required to have degree less than or equal to $x^{2r_k - 1}$ and to be divisible by $x^{2(t_k - r_l)}$. The $p_{ei} \times p_{ei}$ upper triangular polynomial matrix $V_e(x) = [v_{kl}^e(x)]$ is obtained by solving the polynomial matrix identity

$$V_e(x) U_e(x) = I_{p_{ei}} \bmod [x^{2r_{k,l}}], \tag{4.73}$$

where the modulo operation in (4.73) indicates that the (k, l)th polynomial entries on both sides of this identity are equal modulo $x^{2r_{k,l}}$. The coefficients of $V_e(x)$ can be determined by identifying coefficients of increasing orders on both sides of (4.73).

Then, taking into account the parametrization (4.57a)–(4.57b) for the cross block G_{eo}, it is easy to verify that $\tilde{Q}_i$ commutes with E_i and A_i. This completes the first stage of the factorization of G_i.

Step 2: All that remains to do is to factor the skew-symmetric block F_o. Note that since G_i is invertible, $\tilde{G}_i$ and its F_o block are also invertible. F_o obeys the same equation as G_o, so that in the partition $F_o = [F_{kl}^o]$, the subblocks of

$$F_{kl}^o = \begin{bmatrix} G_{kl}^{o1} & F_{kl}^{o2} \\ F_{kl}^{o3} & F_{kl}^{o4} \end{bmatrix} \tag{4.74}$$

admit exactly the same parametrization as the blocks G_{kl}^{oj}. The only difference is that the polynomials $p_{kl}^{oj}(x)$ appearing in (4.61a)–(4.61d) are replaced by polynomials $m_{kl}^{oj}(x)$. The polynomial $m_{kl}^{oj}(x)$ has degree less or equal to $x^{2r'_{k,l}}$, with $r'_{k,l} = \max(r'_k, r'_l)$, and is divisible by $x^{2|r'_k - r'_l|}$. The skew-symmetry of F_o implies also that the polynomials $m_{kl}^{oj}(x)$ satisfy

$$m_{kl}^{o1}(x) = -m_{lk}^{o1}(-x), \tag{4.75a}$$

$$m_{kl}^{o2}(x) = -m_{lk}^{o3}(-x), \tag{4.75b}$$

$$m_{kl}^{o4}(x) = -m_{lk}^{o4}(-x). \tag{4.75c}$$

and that for $k = l$, $m_{kk}^{o1}(x)$ and $m_{kk}^{o4}(x)$ are odd polynomials.

The next step is to construct a matrix Q_o such that the factorization

$$F_o = Q_o^T K_o Q_o \tag{4.76a}$$

holds, where

$$K_o := \bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} 0 & -I_{2r_k'+1} \\ I_{2r_k'+1} & 0 \end{bmatrix} \tag{4.76b}$$

is the odd block of K_i in (2.5c). Without loss of generality, we can assume that the blocks F_{kl}^{oj} have decreasing sizes, so that $2r_1' + 1 \geqslant 2r_2' + 1 \geqslant \cdots \geqslant 2r_{p_{oi}}' + 1$. Partitioning Q_o as $Q_o = [Q_{kl}^o]$, we again require Q_o to be *block upper triangular*, so that $Q_{kl}^o = 0$ for $k > l$ and

$$Q_{kl}^o = \begin{bmatrix} Q_{kl}^{o1} & Q_{kl}^{o2} \\ Q_{kl}^{o3} & Q_{kl}^{o4} \end{bmatrix} \tag{4.77}$$

for $l \geqslant k$, where the blocks Q_{kl}^{oj} admit the parametrization (4.60a)–(4.60d), except that the polynomials $p_{kl}^{oj}(x)$ are replaced by polynomials $u_{kl}^{oj}(x)$. For $k \leqslant l$, the polynomials $u_{kl}^{oj}(x)$ are required to have degree less or equal to $2r_k'$ and to be divisible by $x^{2(r_k'-r_l')}$.

To evaluate the polynomials $u_{kl}^{oj}(x)$ so that the factorization (4.76a) holds, it is convenient to rewrite this identity as a partial parasymmetric rational matrix factorization. Thus, let $M_{kl}^o(x)$ and $U_{kl}^o(x)$ be the 2×2 matrix polynomial blocks given by

$$M_{kl}^o(x) = \begin{bmatrix} m_{kl}^{o1}(x) & m_{kl}^{o2}(x) \\ m_{kl}^{o3}(x) & m_{kl}^{o4}(x) \end{bmatrix}, \qquad U_{kl}^o(x) = \begin{bmatrix} u_{kl}^{o1}(x) & u_{kl}^{o2}(x) \\ u_{kl}^{o3}(x) & u_{kl}^{o4}(x) \end{bmatrix}, \tag{4.78}$$

and consider the $2p_{oi} \times 2p_{oi}$ rational matrix

$$\Phi_o(x) = \left[\frac{M_{kl}^o(x)}{x^{2r'_{k,l}} + 1} \right] \tag{4.79}$$

and the $2p_{oi} \times 2p_{oi}$ polynomial matrix $U_o(x) = [U_{kl}^o(x)]$. The properties (4.75a)–(4.75c) of the polynomials $m_{kl}^{oj}(x)$ imply that $\Phi_o(x)$ is *parasymmetric* i.e.,

$$\Phi_o(x) = (\Phi_o(-x))^T. \tag{4.80}$$

Then, by substituting the parametrization of the blocks Q_{kl}^{oj} inside the factorization (4.76a), we can rewrite it as the partial rational matrix factorization

$$\Phi_o(x) = \pi_-\left\{(U_o(-x))^T S_o(x) U_o(x)\right\}, \tag{4.81}$$

with

$$S_o(x) = \operatorname{diag}\left\{ \frac{1}{x^{2r'_k+1}};\, 1 \leqslant k \leqslant p_{oi} \right\} \otimes \Sigma_2. \tag{4.82}$$

The existence of such a factorization is a consequence of the parasymmetry of $\Phi_o(x)$. First, by adapting the argument of Appendix B, note that because the block F_o is invertible, when $\Phi_o(x)$ is represented as

$$\Phi_o(x) = D_o^{-1}(x) N_o(x) = (N_o(-x))^T D_o^{-1}(x) \tag{4.83a}$$

with

$$D_o(x) = \operatorname{diag}\left\{x^{2r'_k+1};\, 1 \leqslant k \leqslant p_{oi}\right\} \otimes \Sigma_2, \tag{4.83b}$$

the constant matrix $N_o(0)$ is invertible, so that (4.83a) yields both left- and right-coprime matrix function descriptions of $\Phi_o(x)$. This implies that $U_o(0)$ is invertible, and thus the constant 2×2 matrix coefficients $U_{kk,0}^o$ of the diagonal blocks $U_{kk}^o(x)$ are invertible for all k. Then the factorization (4.81) can be constructed block row by block row. Specifically, for $k \leqslant l$, let

$$\tilde{M}_{kl}^{o}(x) := M_{kl}^{o}(x) - \sum_{h=1}^{k-1} x^{2(r_k' - r_h')}\left(U_{hk}^{o}(-x)\right)^T \Sigma_2 U_{hl}^{o}(x). \quad (4.84)$$

Note that $\tilde{M}_{kk}^{o}(x) = -(\tilde{M}_{kk}^{o}(-x))^T$, so that $\tilde{M}_{kk}^{o}(x)$ is para-skew-symmetric. Also, if $M_{kl}^{o}(x)$ is divisible by $x^{2(r_k' - r_l')}$ and $U_{hl}^{o}(x)$ is divisible by $x^{2(r_h' - r_l')}$ for $h < l$, then $\tilde{M}_{kl}^{o}(x)$ is also divisible by $x^{2(r_k' - r_l')}$. Then, for $1 \leqslant k \leqslant p_i$ and $k \leqslant l \leqslant p_i$, the factorization (4.81) is achieved by solving the equations

$$U_{kk}^{oT}(-x)\Sigma_2 U_{kk}^{o}(x) = \tilde{M}_{kk}(x) \bmod x^{2r_k'+1}, \quad (4.85a)$$

$$U_{kk}^{oT}(-x)\Sigma_2 U_{kl}^{o}(x) = \tilde{M}_{kl}(x) \bmod x^{2r_k'+1}. \quad (4.85b)$$

The factorization (4.85a) is obtained by matching the 2×2 matrix coefficients of successive orders on both sides of (4.85a). Specifically, consider the polynomial expansions

$$\tilde{M}_{kk}^{o}(x) = \sum_{s=0}^{2r_k'} \tilde{M}_{kk,s}^{o} x^s, \qquad U_{kk}^{o}(x) = \sum_{s=0}^{2r_k'} U_{kk,s}^{o} x^s. \quad (4.86)$$

Because of the para-skew-symmetry of $\tilde{M}_{kk}^{o}(x)$, its constant coefficient

$$\tilde{M}_{kk,0}^{o} = \begin{bmatrix} 0 & \tilde{m}_{kk,0}^{o2} \\ -\tilde{m}_{kk,0}^{o2} & 0 \end{bmatrix} \quad (4.87)$$

is a skew-symmetric matrix, where $\tilde{m}_{kk,0}^{o2}$ is nonzero, since $\tilde{M}_{kk,0}^{o}$ is invertible. Then the factorization

$$U_{kk,0}^{oT}\Sigma_2 U_{kk,0}^{o} = \tilde{M}_{kk,0}^{o} \quad (4.88a)$$

holds with

$$U_{kk,0}^{o} = \begin{bmatrix} -\tilde{m}_{kk,0}^{o2} & 0 \\ 0 & 1 \end{bmatrix}. \quad (4.88b)$$

Furthermore, the equation obtained by matching the coefficients of x^s on both sides of (4.85a) takes the form

$$U_{kk,0}^{ot}\Sigma_2 U_{kk,s}^{o} + (-1)^s U_{kk,s}^{oT}\Sigma_2 U_{kk,0}^{o} = C_{kk,s} \tag{4.89}$$

where the 2×2 matrix $C_{kk,s}$ satisfies $C_{kk,s} = (-1)^{s+1}C_{kk,s}^T$. Taking into account the form (4.85b) of $U_{kk,0}^o$, it is easy to verify that this equation admits a solution $U_{kk,s}^o$. Substituting the matrix polynomial $U_{kk}^{oT}(-x)$ inside (4.85b), we can then match the coefficients of increasing powers of x on both sides of (4.85b) to evaluate the matrix coefficients of $U_{kl}^o(x)$ for $l > k$. These coefficients are uniquely determined because $U_{kk,0}^o$ is invertible. The invertibility of $U_{kk,0}^o$ also ensures that since $\tilde{M}_{kl}^o(x)$ is divisible by $x^{2(r_k^i - r_l^i)}$, so is $U_{kl}^o(x)$.

This construction yields the factorization (4.76a), where the parametrization (4.60a)–(4.60d) of the blocks Q_{kl}^{oj} implies that Q_o commutes with both E_o and A_o. Then, if we define

$$Q_i := \begin{bmatrix} I & 0 \\ 0 & Q_o \end{bmatrix} \tilde{Q}_i, \tag{4.90}$$

Q_i commutes with E_i and A_i, and after combining (4.71) and (4.76a), we obtain the factorization (4.12) of G_i.

4.3. *Conclusion*

Having now constructed factorizations of the form (4.12) for blocks of each type, we employ the resulting factors to modify the matrices V and W of (4.1) to ensure (4.2) holds. If we denote

$$\overline{V} := \bigoplus_{i=1}^{l} \overline{V}_i, \qquad \overline{W} := \bigoplus_{i=1}^{l} \overline{W}_i, \tag{4.91}$$

where $\overline{V}_i$ and $\overline{W}_i$ are given by (4.13), the commutation relations (4.14) can be written compactly as

$$\left(\bigoplus_{i=1}^{l} (sE_i - tA_i) \right) \overline{V} = \overline{W} \left(\bigoplus_{i=1}^{l} (sE_i - tA_i) \right). \tag{4.92}$$

Mutliplying Equation (4.1) on the right by $\overline{V}$ gives

$$(sE - tA)V\overline{V} = W\overline{W}\left(\bigoplus_{i=1}^{l}(sE_i - tA_i)\right), \tag{4.93}$$

or equivalently

$$(sE - tA)\hat{V} = \hat{W}\left(\bigoplus_{i=1}^{l}(sE_i - tA_i)\right), \tag{4.94}$$

with $\hat{V} := V\overline{V}$ and $\hat{W} := W\overline{W}$.

Finally, we find

$$\begin{aligned}\hat{W}^T K\hat{W} &= \overline{W}^T W^T KW\overline{W}\\ &= \overline{W}^T\left(\bigoplus_{i=1}^{l} G_i\right)\overline{W}\\ &= \left(\bigoplus_{i=1}^{l} Q_i^{-T}\right)\left(\bigoplus_{i=1}^{l} G_i\right)\left(\bigoplus_{i=1}^{l} Q_i^{-1}\right)\\ &= \bigoplus_{i=1}^{l} K_i,\end{aligned} \tag{4.95}$$

where we have used the identities (4.10) and (4.12), and where, depending on whether they are of type 1, 2, 3, or 4, the skew-symmetric blocks K_i have the structure (2.2c), (2.3c), (2.4c), or (2.5c). Thus, the real matrices $\hat{W}$ and $\hat{V}$ obey the identities (4.1) and (4.2) of Theorem 2.1.

It remains to show that to each symplectic pair $(P(s,t), K)$, with K given by (1.1), corresponds a unique canonical form $(\tilde{P}(s,t), \tilde{K})$. Let V_1, W_1, V_2, and W_2 be four nonsingular matrices such that

$$(sE - tA)V_m = W_m\tilde{P}_m(s,t), \quad W_m^T KW_m = \tilde{K}_m, \qquad m = 1,2, \tag{4.96}$$

where the pairs $(\tilde{P}_m(t,s), \tilde{K}_m)$, with $m = 1, 2$, have the structure described in Theorem 2.1. It is obvious from the previous construction that $\tilde{P}_1(s,t) = \tilde{P}(s,t)$ and that $\tilde{K}_1$ and $\tilde{K}_2$ may differ only by the signs $\kappa_k = \pm 1$ appearing in the blocks of type 3 and 4. We now show that these signs need actually to be the same, so that $\tilde{K}_1 = \tilde{K}_2$. Denoting

$$\tilde{P}(s,t) = \tilde{P}_1(s,t) = \tilde{P}_2(s,t), \tag{4.97}$$

the equations (4.96) yield

$$\tilde{W}\tilde{P}(s,t) = \tilde{P}(s,t)\tilde{V}, \tag{4.98a}$$

$$\tilde{K}_1 = \tilde{W}^T\tilde{K}_2\tilde{W}, \tag{4.98b}$$

where $\tilde{W} := W_2^{-1}W_1$ and $\tilde{V} := V_2^{-1}V_1$. The identity (4.98a) implies that $\tilde{V}$ and $\tilde{W}$ have a block-diagonal structure matching the structure of $\tilde{K}_1 = \bigoplus_{i=1}^{l} K_{1i}$ and $\tilde{K}_2 = \bigoplus_{i=1}^{l} K_{2i}$:

$$\tilde{V} = \bigoplus_{i=1}^{l} \tilde{V}_i, \qquad \tilde{W} = \bigoplus_{i=1}^{l} \tilde{W}_i. \tag{4.99}$$

Then, Equations (4.98a)–(4.98b) may be written blockwise as

$$\tilde{W}_i(sE_i - tA_i) = (sE_i - tA_i)\tilde{V}_i, \tag{4.100a}$$

$$K_{1i} = \tilde{W}_i^T K_{2i}\tilde{W}_i, \tag{4.100b}$$

with $i = 1, 2, \ldots, l$. In particular, let $\bar{i}$ be the index corresponding to a block of type 3 (the corresponding $E_{\bar{i}}$ is nonsingular) and $S := A_{\bar{i}}E_{\bar{i}}^{-1}$. The corresponding $K_{1\bar{i}}$ and $K_{2\bar{i}}$ blocks have the structure

$$K_{1\bar{i}} = \bigoplus_{k=1}^{p_{\bar{i}}} \kappa_{1k}\left(\Sigma_{r_k} \otimes \Sigma_2^{r_k}\right), \qquad K_{2\bar{i}} = \bigoplus_{k=1}^{p_{\bar{i}}} \kappa_{2k}\left(\Sigma_{r_k} \otimes \Sigma_2^{r_k}\right). \tag{4.101}$$

From (4.100a) it immediately follows that $\tilde{W}_{\bar{i}}$ commutes with S, which in turn implies that $\tilde{W}_{\bar{i}}$ commutes with $J_{\bar{i}} = \bigoplus_{k=1}^{p_{\bar{i}}} J_{2r_k}(jb_{\bar{i}}, -jb_{\bar{i}})$:

$$\tilde{W}_{\bar{i}}J_{\bar{i}} = J_{\bar{i}}\tilde{W}_{\bar{i}}. \tag{4.102}$$

But it is shown in [6] that if $K_{1\bar{i}}$ and $K_{2\bar{i}}$ are given by (4.101) and $\kappa_{1k} \neq \kappa_{2k}$ for at least one value of k, then the Hamiltonian normal forms $(J_{\bar{i}}, K_{1\bar{i}})$ and

(J_i, K_{2i}) cannot satisfy the equivalence relation (4.100b), (4.102). Therefore, $K_{1i} = K_{2i}$ (up to a permutation of blocks of the same size). A similar argument holds for blocks of type 4.

APPENDIX A. INVERTIBILITY OF G_e and G_o

In this appendix, we prove that in the partition (4.49) of a matrix G_i corresponding to a block of type 4, the even and odd blocks G_e and G_o are both invertible. The proof relies on a property of determinants of block matrices with upper triangular Toeplitz blocks.

Consider a matrix $T = [T_{kl}]$ whose blocks T_{kl} have size $r_k \times r_l$ with $1 \leqslant k, l \leqslant p$. For $r_k \geqslant r_l$, T_{kl} admits the parametrization

$$T_{kl} = t_{kl}(Z_{r_k}) \begin{bmatrix} 0_{(r_k - r_l)\times r_l} \\ I_{r_l} \end{bmatrix}, \tag{A.1a}$$

where $t_{kl}(x)$ is a polynomial of degree less or equal to $r_k - 1$ which is divisible by $x^{r_k - r_l}$. Similarly, for $r_l \geqslant r_k$ we have

$$T_{kl} = \begin{bmatrix} I_{r_k} & 0_{(r_l - r_k)\times r_k} \end{bmatrix} t_{kl}(Z_{r_l}), \tag{A.1b}$$

where $t_{kl}(x)$ is a polynomial of degree less or equal to $r_l - 1$ which is divisible by $x^{r_l - r_k}$. We assume that the diagonal blocks have decreasing size, and that there are n_1 diagonal blocks of size s_1, n_2 blocks of size $s_2, \dots,$ and n_q blocks of size s_q with $s_1 > s_2 > \cdots > s_q$. Let

$$C_h = \left[t_{kl}(0);\ \sum_{j=1}^{h-1} n_j + 1 \leqslant k, l \leqslant \sum_{j=1}^{h} n_j \right] \tag{A.2}$$

for $1 \leqslant h \leqslant q$. The matrix C_h is formed by the constant coefficients of the polynomials $t_{kl}(x)$ associated to all the square blocks of T with size $s_h \times s_h$. Because there are n_h diagonal blocks of size s_h, the matrix C_h has size $n_h \times n_h$. Then, the determinant of T can be evaluated as follows.

LEMMA A.1. *Let T be a block matrix whose blocks have the upper triangular Toeplitz structure* (A.1a)–(A.1b). *Then*

$$\det T = \prod_{h=1}^{q} (\det C_h)^{s_h}. \tag{A.3}$$

Proof. Let $\delta_1 = s_1 - s_2$, $\delta_2 = s_2 - s_3, \dots, \delta_{q-1} - s_q$. Then we perform a permutation of the rows and columns of T where for $1 \leqslant j \leqslant \delta_1$, we select the rows and columns of T corresponding to the jth rows and columns of the blocks of size s_1. Next, for $1 \leqslant j \leqslant \delta_2$, we pick the rows and columns of T corresponding to the $\delta_1 + j$th rows and columns of the blocks of size s_1 and the jth rows and columns of the blocks of size s_2. Continuing this process, we finally select for $1 \leqslant j \leqslant s_q$ the rows and columns of T corresponding to the $\Sigma_{i=1}^{q-1}\delta_i + j$th rows and columns of block size s_1, the $\Sigma_{i=2}^{q-1}\delta_i + j$th rows and columns of blocks of size s_1, the $\Sigma_{i=2}^{q-1}\delta_i + j$th rows and columns of blocks of size $s_2, \dots,$ and the jth rows and columns of the blocks of size s_q. This yields a matrix $\hat{T}$ which is block upper triangular, whose diagonal blocks are block lower triangular, and where the diagonal blocks are first δ_1 copies of C_1, followed by δ_2 copies of the pair (C_1, C_2), and then s_q copies of the q-tuple $(C_1, C_2, \dots, C_q)$. As a consequence, $\det T = \det \hat{T}$ has the form (A.3).

Example. To illustrate the technique described above to evaluate $\det T$, consider the matrix

$$T = \left[\begin{array}{ccc|ccc|cc|cc} \alpha & \beta & \gamma & \delta & \epsilon & \zeta & i & j & k & l \\ 0 & \alpha & \beta & 0 & \delta & \epsilon & 0 & i & 0 & k \\ 0 & 0 & \alpha & 0 & 0 & \delta & 0 & 0 & 0 & 0 \\ \hline \theta & \kappa & \lambda & \mu & \nu & \rho & m & n & p & q \\ 0 & \theta & \kappa & 0 & \mu & \nu & 0 & m & 0 & p \\ 0 & 0 & \theta & 0 & 0 & \mu & 0 & 0 & 0 & 0 \\ \hline 0 & r & s & 0 & t & u & a & b & c & d \\ 0 & 0 & r & 0 & 0 & t & 0 & a & 0 & c \\ \hline 0 & v & w & 0 & x & y & e & f & g & h \\ 0 & 0 & v & 0 & 0 & x & 0 & e & 0 & g \end{array}\right]. \tag{A.4}$$

It has $n_1 = 2$ blocks of size $s_1 = 3$, and $n_2 = 2$ blocks of size $s_2 = 2$, and we have

$$C_1 = \begin{bmatrix} \alpha & \delta \\ \theta & \mu \end{bmatrix}, \qquad C_2 = \begin{bmatrix} a & c \\ e & g \end{bmatrix}. \tag{A.5}$$

By permuting its rows and columns as indicated in the proof of Lemma A.1, we obtain the matrix

$$\hat{T} = \left[\begin{array}{cc|cccc|cccc} \alpha & \delta & \beta & \epsilon & i & k & \gamma & \zeta & j & l \\ \theta & \mu & \kappa & \nu & m & p & \lambda & \rho & n & q \\ \hline 0 & 0 & \alpha & \delta & 0 & 0 & \beta & \epsilon & i & k \\ 0 & 0 & \theta & \mu & 0 & 0 & \kappa & \nu & m & p \\ 0 & 0 & r & t & a & c & s & u & b & d \\ 0 & 0 & v & x & e & g & w & y & f & h \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \alpha & \delta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \theta & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & r & t & a & c \\ 0 & 0 & 0 & 0 & 0 & 0 & v & x & e & g \end{array}\right] \tag{A.6}$$

from which it is then easy to verify that

$$\det T = \det \hat{T} = (\det C_1)^3 (\det C_2)^2. \tag{A.7}$$

An important feature of the expression (A.3) for the determinant of T is that it does not depend on the coefficients of the blocks T_{kl} with different row and column sizes, i.e. with $r_k \neq r_l$, so that when evaluating the determinant of T, these blocks can be set equal to zero. Consider now the matrix G_i with the structure (4.49), where we assume that the blocks of the even and odd matrices G_e and G_o have decreasing sizes. Let

$$L_i = \left(\bigoplus_{k=1}^{p_{ei}} \Sigma_{2r_k} \right) \oplus \left(\bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} \Sigma_{2r_l'+1} & 0 \\ 0 & I_{2r_l'+1} \end{bmatrix} \right), \tag{A.8a}$$

$$R_i = \left(\bigoplus_{k=1}^{p_{ei}} I_{2r_k} \right) \oplus \left(\bigoplus_{k=1}^{p_{oi}} \begin{bmatrix} I_{2r_l'+1} & 0 \\ 0 & \Sigma_{2r_l'+1} \end{bmatrix} \right). \tag{A.8b}$$

Then, by taking into account the structure (4.52a)–(4.52b), (4.57a)–(4.58b), and (4.60a)–(4.61d) of the even, cross, and odd blocks of G_i, we find that G_i can be expressed as

$$G_i = L_i T_i R_i \tag{A.9a}$$

with

$$T_i = \begin{bmatrix} T_e & T_{eo} \\ T_{oe} & T_o \end{bmatrix}, \tag{A.9b}$$

and where T_e, T_{eo}, and T_o can be partitioned into subblocks which have all the upper triangular Toeplitz structure (A.1a)–(A.1b). Because the subblocks of T_{eo} have an even number of rows and an odd number of columns, they do not contribute to the determinant of T_i, so that

$$\det T_i = \det T_e \det T_o. \tag{A.10}$$

Observing that $\det \Sigma_r = 1$ for all r, we find

$$\det G_i = \det G_e \det G_o, \tag{A.11}$$

so that G_i is invertible if and only if G_e and G_o are both invertible.

APPENDIX B. INVERTIBILITY OF $N_i(0)$

In this section, we prove that for blocks of type 3, the invertibility of G_i implies that in the representation (4.43a) of $\Phi_i(x)$, the matrix $N_i(0)$ is invertible.

First note that the unitary matrix

$$U_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -j & j \end{bmatrix} \tag{B.1}$$

diagonalizes Σ_2, since

$$U_2^H \Sigma_2 U_2 = \operatorname{diag}\{j, -j\}. \tag{B.2}$$

Let P_{eo} be the permutation which for a matrix of even size permutes all the rows so that the rows with an odd index are placed ahead of those with an even index, while preserving the internal order of the odd and even index sets. Let also

$$L_i = \left(\bigoplus_{k=1}^{p_i} \left(\Sigma_{r_k} \otimes \Sigma_2^{r_k} U_2 \right) \right) P_{eo}, \tag{B.3a}$$

$$R_i = P_{eo}^T \left(\bigoplus_{k=1}^{p_i} \left(I_{r_k} \otimes U_2^H \right) \right). \tag{B.3b}$$

Then, by taking into account the parametrization (4.28b) of the blocks G_{kl} of G_i, where Γ_{kl}^R and Γ_{kl}^I admit the structure (4.31a)–(4.31b), it is easy to verify that

$$G_i = L_i (T_i \oplus T_i^*) R_i, \tag{B.4}$$

where the blocks of $T_i = [T_{kl}]$ have the upper triangular Toeplitz structure (A.1a)–(A.1b) with the polynomials $t_{kl}(x)$ replaced by the complex polynomials $p_{kl}(x)$ given by (4.37a). For the matrix T_i, let C_h be the square matrices defined by (A.2). By applying Lemma A.1, we find

$$\det G_i = \prod_{h=1}^{q} |\det C_h|^{2 s_h}. \tag{B.5}$$

However, by taking into account the definition (4.38) of $\Phi_i(x)$ and the expression (A.2) for C_h, it is easy to check that the matrix $N_i(0)$ is block lower triangular, and its diagonal blocks are C_h for $1 \leqslant h \leqslant q$. This implies

$$\det N_i(0) = \prod_{h=1}^{q} \det C_h, \tag{B.6}$$

and comparing (B.5) and (B.6), we see that $N_i(0)$ is invertible if and only if G_i is invertible.

REFERENCES

1 N. Burgoyne and R. Cushman, Normal forms for real Hamiltonian systems, in *The 1976 Ames Research Center (NADA) Conference on Geometric Control Theory* (C. Martin and R. Hermann, eds.), Math Sci. Press, Brookline, Mass., 1977, pp. 483–529.

2 A. Ciampi, *Classical Hamiltonian Linear Systems*, Queen's Papers in Pure and Appl. Math. 31, Queen's Univ., Kingston, Ontario, Canada, 1972.
3 A. Ciampi, Classification of Hamiltonian linear systems, *Indiana Univ. Math. J.* 23:513–526 (1973).
4 I. K. Cikunov, On a class of isometric transformations of a symplectic or orthogonal vector space (in Russian), *Ukrain. Mat. Zh.* 18:122–127 (1966).
5 I. K. Cikunov, On the structure of isometric transformations of a sympectic or orthogonal vector space (in Russian), *Ukrain. Mat. Zh.* 18:79–93 (1966).
6 D. Z Djokovic, J. Patera, P. Winternitz, and H. Zassenhaus, Normal forms of elements of classical real and complex Lie and Jordan algebras, *J. Math. Phys.* 24:1363–1374 (June 1983).
7 F. R. Gantmacher, *The Theory of Matrices*, Vol. 2, Chelsea, 1959.
8 I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products*, Birkhäuser, Basel, 1983.
9 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U.P., 1985.
10 R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge U.P., 1991.
11 T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N.J., 1980.
12 P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Oxford U.P., New York, 1995.
13 A. J. Laub and K. Meyer, Canonical forms for symplectic and Hamiltonian matrices, *Celestial Mech.* 9:213–238 (1974).
14 B. C. Levy, Characterization of multivariate Gaussian reciprocal diffusions, *J. Multivariate Anal.*, to appear.
15 G. W. Stewart, On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$, *J. Numer. Anal.* 9:669–686 (1972).
16 H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices*, Blackie, Loneon, 1932.
17 J. Williamson, On the algebraic problem concerning the normal forms of linear dynamic systems, *Amer. J. Math.* 58:141–163 (1936).
18 J. Williamson, On the normal forms of linear canonical transformations in dynamics, *Amer. J. Math.* 59:599–617 (1937).
19 J. Williamson, The exponential representation of canonical matrices, *Amer. J. Math.* 61:897–911 (1939).
20 H. K. Wimmer, Normal forms of symplectic pencils and the discrete-time algebraic Riccati equation, *Linear Algebra Appl.* 147:411–440 (1991).

Received 4 November 1996; final manuscript accepted July 1997